LAGRANGIAN AND HAMILTONIAN FORMALISM FOR CONSTRAINED VARIATIONAL PROBLEMS

PAOLO PICCIONE AND DANIEL V. TAUSK

ABSTRACT. We consider solutions of Lagrangian variational problems with linear constraints on the derivative. More precisely, given a smooth distribution $\mathcal{D} \subset TM$ on M and a time-dependent Lagrangian L defined on $\mathcal{D}$, we consider an action functional $\mathcal{L}$ defined on the set $\Omega_{PQ}(M,\mathcal{D})$ of horizontal curves in M connecting two fixed submanifolds $P, Q \subset M$. Under suitable assumptions, the set $\Omega_{PQ}(M,\mathcal{D})$ has the structure of a smooth Banach manifold and we can thus study the critical points of $\mathcal{L}$. If the Lagrangian L satisfies an appropriate *hyper-regularity* condition, we associate to it a degenerate Hamiltonian H on TM^* using a general notion of *Legendre transform* for maps on vector bundles. We prove that the solutions of the Hamilton equations of H are precisely the critical points of $\mathcal{L}$. In the particular case where L is given by the quadratic form corresponding to a positive definite metric on $\mathcal{D}$, we obtain the well-known characterization of the *normal geodesics* in sub-Riemannian geometry (see [10]); by adding a potential energy term to L, we reobtain the equations of motion for the *Vakonomic mechanics* with non holonomic constraints (see [8]).

1. INTRODUCTION

The aim of this paper is to generalize to the context of constrained variational problems some classical results about the correspondence between Lagrangian and Hamiltonian formalisms (see for instance [1]). Particular cases of this theory are the *sub-Riemannian geodesic problem* (see for instance [10, 11, 13, 17]), and the so called *Vakonomic* approach to the non holonomic mechanics (see for instance [2, 5, 8, 19]).

The constrained variational problem studied is modelled by the following setup: we consider an n-dimensional differentiable manifold M endowed with a smooth distribution $\mathcal{D} \subset TM$ of rank k; moreover, we assume that it is given a (possibly time-dependent) Lagrangian function L on $\mathcal{D}$. In the non holonomic mechanics, M represents the configuration space, $\mathcal{D}$ the constraint, and L is typically the difference between the kinetic and a potential energy. In the sub-Riemannian geodesic problem, L is simply the quadratic form corresponding to a positive definite metric on $\mathcal{D}$.

The solutions of the constrained variational problem are given by curves $\gamma : [a,b] \to M$ that are critical points of the action functional $\mathcal{L}(\gamma) = \int_a^b L\big(t, \gamma(t), \dot\gamma(t)\big)\,\mathrm{d}t$ defined on the space:

$$\Omega_{PQ}\big([a,b], M, \mathcal{D}\big) = \big\{\gamma : [a,b] \xrightarrow{C^1} M : \gamma(a) \in P,\ \gamma(b) \in Q,\ \gamma'(t) \in \mathcal{D} \text{ for all } t\big\}$$

of horizontal curves of class C^1 in M connecting two fixed submanifolds $P, Q \subset M$. It is well-known that the set $\Omega_{PQ}\big([a,b], M, \mathcal{D}\big)$ is in general not a submanifold of the Banach manifold of C^1 curves $\gamma : [a,b] \to M$; when P and Q are points, the singularities of

Date: September 2001.
2000 *Mathematics Subject Classification.* 37J05, 37J50, 37J60, 53C17, 70H03, 70H05.
The authors are partially sponsored by CNPq, Brazil.

$\Omega_{PQ}\big([a,b],M,\mathcal{D}\big)$ are known in the context of sub-Riemannian geometry as *abnormals extremals* (see [4, 10, 11, 12, 13, 17]). Such singularities can be nicely described using the canonical symplectic structure of the cotangent bundle TM^* (see Corollary 3.7). In this paper we are interested in studying the action functional $\mathcal{L}$ in the *regular part* of $\Omega_{PQ}\big([a,b],M,\mathcal{D}\big)$. We remark that in several important cases the set $\Omega_{PQ}\big([a,b],M,\mathcal{D}\big)$ contains no singular curves (see, for instance, Corollary 3.8 and Remark 3.9).

Recall from [1] that when a Lagrangian function $L : TM \to I\!R$ is *hyper-regular* then the critical points of the corresponding (unconstrained) variational problem are given by the solutions of the Hamilton equations corresponding to a Hamiltonian $H : TM^* \to I\!R$ which corresponds to L by means of the *Legendre transform*. The Legendre transform described in [1] can be generalized in a straightforward way to general vector bundles; namely, if $L : \xi \to I\!R$ is a smooth map on a vector bundle ξ which is hyper-regular (in a suitable sense) then one can naturally associate to it a smooth map $H : \xi^* \to I\!R$ on the dual bundle ξ^*. At such level of generality, the Legendre transform does not seem to have a meaningful interpretation in the context of calculus of variations, as it does in the case $\xi = TM$. Our goal is to show that when $\xi = \mathcal{D}$ is a vector subbundle of a tangent bundle TM (i.e., a distribution on M) then the Legendre transform for smooth maps on $\mathcal{D}$ has a nice application to the study of constrained variational problems. The key observation here is that, when passing to the dual bundles, the *inclusion* arrow $\mathcal{D} \to TM$ reverses and gives rise to a *projection* arrow $TM^* \to \mathcal{D}^*$; thus, while a *constrained Lagrangian* $L : \mathcal{D} \to I\!R$ has no canonical extension to a Lagrangian on TM, its Legendre transform $H_0 : \mathcal{D}^* \to I\!R$ naturally induces a map $H : TM^* \to I\!R$ given by the composition of H_0 and the projection $TM^* \to \mathcal{D}^*$. Our main result (Theorem 4.1) is that the critical points of the constrained action functional $\mathcal{L}$ are the solutions of the Hamilton equations of H satisfying suitable boundary conditions. Observe that, unless $\mathcal{D} = TM$, the Hamiltonian H is *always degenerate* and thus it cannot arise as the Legendre transform of a hyper-regular Lagrangian on the whole tangent bundle TM.

In the particular case where P and Q are single points of M, $\mathcal{D}$ is endowed with a smoothly varying positive definite inner product g and L is given by $L(t,q,\dot q) = \frac{1}{2}g(\dot q,\dot q)$, then the solutions of the corresponding Hamiltonian H are known in the context of *sub-Riemannian geometry* as the *normal extremals* of $(M,\mathcal{D},g)$. The critical points of the constraint defining $\Omega_{PQ}\big([a,b],M,\mathcal{D}\big)$ are the abnormal extremals. In particular, we obtain a variational proof of [10, Theorem 1]. By adding a potential energy term to L, the Hamilton equations of H become the equations of motion for the Vakonomic mechanics (see [8]). Theorem 4.1 thus provides a unifying approach for the study of Lagrangian variational problems with linear constraints in the derivative; it also provides the appropriate setting for the study of the *second variation* of a constrained Lagrangian action functional and for the development of an *index theory* for such functional using the notion of *Maslov index* for a solution of a Hamiltonian (see [18]).

The proof of Theorem 4.1 is based on the method of Lagrangian multipliers, which is used to pass from a constrained Lagrangian variational problem to a non constrained one. The main technical difficulty is the proof of the regularity of the Lagrangian multiplier (Lemma 4.9); such proof is based on a suitable version of Schwartz's generalized functions calculus which is developed in Subsection 4.1.

We give a brief description of the material presented in each section of the paper.

In Subsection 2.1 we describe a general notion of Legendre transform. In Subsection 2.2 we recall some standard results concerning the correspondence between hyper-regular Lagrangians and Hamiltonians and in Section 3 we present some well-known facts about the

manifold structure of the set of horizontal curves connecting two fixed submanifolds of a given manifold.

In Section 4 we state the main result of the paper (Theorem 4.1), that establishes the correspondence between the critical points of the action functional of a hyper-regular constrained Lagrangian and the solutions of the corresponding degenerate Hamiltonian. The proof of Theorem 4.1 is given in Subsection 4.2. In Subsection 4.1 it is presented a suitable version of Schwartz's generalized functions calculus, needed for technical reasons in the proof of Theorem 4.1.

2. THE LEGENDRE TRANSFORM. LAGRANGIANS AND HAMILTONIANS ON MANIFOLDS

In this section we recall some classical results from [1] which are presented in a more general context needed for the statement and the proof of Theorem 4.1. In Subsection 2.1 we present a general version of the Legendre transform for vector spaces; we then apply it fiberwise to obtain a notion of Legendre transform for fiber bundles. In Subsection 2.2 we present the classical Hamiltonian formulation for the variational problem corresponding to a hyper-regular (non constrained) Lagrangian. The standard results from [1] are proven in a slightly more general setup; namely, we consider curves with endpoints varying in submanifolds, time-dependent Lagrangians and rather weak regularity assumptions for the data.

2.1. *The Legendre transform*

Let ξ_0 be a real finite-dimensional vector space, let ξ_0^* denote its dual, and let $Z : U \to I\!R$ be a function of class C^2 defined on an open subset $U \subset \xi_0$.

Definition 2.1. Assume that the differential $\mathrm{d}Z$ is a diffeomorphism onto an open subset $V \subset \xi_0^*$. The *Legendre transform* of Z is the C^1 map $Z^* : V \to I\!R$ defined by:

$$Z^* = E_Z \circ (\mathrm{d}Z)^{-1}, \tag{2.1}$$

where $E_Z : U \to I\!R$ is given by

$$E_Z(v) = \mathrm{d}Z(v)\, v - Z(v), \quad v \in U. \tag{2.2}$$

Lemma 2.2. *Using the canonical identification of ξ_0 and its bi-dual ξ_0^{**}, the map $\mathrm{d}Z^*$ is the inverse of $\mathrm{d}Z$. Therefore, Z^* is a map of class C^2.*

Proof. Differentiating the equality $Z^* \circ \mathrm{d}Z = E_Z$ and (2.2), we obtain:

$$\mathrm{d}Z^*\big(\mathrm{d}Z(v)\big) \circ \mathrm{d}^2 Z(v) = \mathrm{d}E_Z(v), \quad \mathrm{d}E_Z(v) = \hat{v} \circ \mathrm{d}^2 Z(v),$$

where $\hat{v} \in \xi_0^{**}$ denotes evaluation at v. Since $\mathrm{d}^2 Z(v) : \xi_0 \to \xi_0^*$ is an isomorphism, the conclusion follows. $\square$

Corollary 2.3. $Z^{**} = Z$.

Proof. By Lemma 2.2, we have:

$$Z^{**} = E_{Z^*} \circ (\mathrm{d}Z^*)^{-1} = E_{Z^*} \circ \mathrm{d}Z.$$

Hence, by definition of E_{Z^*}, we get:

$$\begin{aligned} E_{Z^*}\big(\mathrm{d}Z(v)\big) &= \mathrm{d}Z^*\big(\mathrm{d}Z(v)\big)\,\mathrm{d}Z(v) - Z^*\big(\mathrm{d}Z(v)\big) = \\ &= \mathrm{d}Z(v)\, v - E_Z(v) = Z(v). \quad \square \end{aligned}$$

Let now M be a smooth manifold and $\pi : \xi \to M$ be a smooth vector bundle over M; for $m \in M$, we denote by ξ_m the fiber $\pi^{-1}(m)$. The dual bundle of ξ will be denoted by ξ^*; the bi-dual ξ^{**} is canonically identified with ξ.

Let $Z : U \subset \xi \to I\!R$ be a map such that, for every $m \in M$, $U \cap \xi_m$ is open in ξ_m and the restriction of Z to $U \cap \xi_m$ is of class C^2.

Definition 2.4. The *fiber derivative* $\mathbb{F}Z : U \to \xi^*$ is the map defined by:

$$\mathbb{F}Z(v) = \mathrm{d}(Z|_{U\cap\xi_m})(v), \quad v \in U, \tag{2.3}$$

where $m = \pi(v)$. Let $V \subset \xi^*$ denote the image of $\mathbb{F}Z$. We say that Z is *regular* if for each $m \in M$, the set $V \cap \xi_m$ is open in ξ_m and the restriction of $\mathbb{F}Z$ to $U \cap \xi_m$ is a local diffeomorphism; Z is said to be *hyper-regular* if for each m such restriction is a diffeomorphism onto $V \cap \xi^*_m$. If Z is hyper-regular, we define the *Legendre transform* of Z as the map $Z^* : V \to I\!R$ whose restriction to $V \cap \xi_m$ is the Legendre transform of the restriction of Z to $U \cap \xi_m$.

In analogy with (2.2) we also set:

$$E_Z(v) = \mathbb{F}Z(v)\, v - Z(v), \quad v \in U; \tag{2.4}$$

obviously $Z^* = E_Z \circ \mathbb{F}Z^{-1}$.

Applying Lemma 2.2 and Corollary 2.3 fiberwise, we obtain immediately the following:

Proposition 2.5. *Assume that $Z : U \subset \xi \to I\!R$ is hyper-regular. Then, for each $m \in M$, the restriction of Z^* to $V \cap \xi^*_m$ is of class C^2. Moreover, $\mathbb{F}Z$ and $\mathbb{F}Z^*$ are mutually inverse bijections and $Z^{**} = Z$.* □

2.2. *Time dependent Lagrangians and Hamiltonians on manifolds*

Let M be a smooth n-dimensional manifold and let TM, TM^* denote respectively the tangent and the cotangent bundle of M; with a slight abuse of notation, we will denote both the projections of TM and of TM^* by π. Consider the following vector bundles:

$$\xi = I\!R \times TM \xrightarrow{\mathrm{Id}\times\pi} I\!R \times M, \quad \xi^* = I\!R \times TM^* \xrightarrow{\mathrm{Id}\times\pi} I\!R \times M.$$

Observe that the fiber $\xi_{(t,m)}$ is $\{t\} \times T_mM$ and that $\xi^*_{(t,m)} = \{t\} \times T_mM^*$.

Definition 2.6. A *(time-dependent) Lagrangian on* M is a function $L : U \subset \xi \to I\!R$ defined on an open set $U \subset \xi$ and satisfying the following regularity conditions:

1. L is continuous;
2. for each $t \in I\!R$, the map $L(t, \cdot)$ is of class C^1 on $U \cap \big(\{t\} \times TM\big)$ and its differential is continuous on U;
3. for each $t \in I\!R$, the map $\mathbb{F}L(t, \cdot) : U \cap \big(\{t\} \times TM\big) \to \{t\} \times TM^*$ is of class C^1.

A *(time-dependent) Hamiltonian on* M is a function $H : V \subset \xi^* \to I\!R$ defined on an open set $V \subset \xi^*$ and satisfying the following regularity conditions:

1. for all $t \in I\!R$, the map $H(t, \cdot)$ is of class C^1 on $V \cap \big(\{t\} \times TM^*\big)$;
2. for each $(t, m) \in I\!R \times M$, the restriction of H to $V \cap \xi^*_{(t,m)}$ is of class C^2.

We use the notions of regularity and hyper-regularity given in Definition 2.4 for Lagrangians and Hamiltonians on manifolds.

Using the Legendre transform defined in Subsection 2.1 (Definition 2.4), given a hyper-regular Lagrangian L on M, the map $H = L^*$ is a hyper-regular Hamiltonian on M.

Namely, the fact that $H(t,\cdot)$ is of class C^1 follows by applying the Inverse Function Theorem to the map $\mathbb{F}L(t,\cdot)$; moreover, the fact that $V=\mathbb{F}L(U)$ is open in ξ^* follows from the Theorem of Invariance of Domain (see [14]) by observing that $\mathbb{F}L$ is continuous and injective[1].

If H is the hyper-regular Hamiltonian obtained by Legendre transform from the Lagrangian L, then by Proposition 2.5, we have that $H^*=L$, and that $\mathbb{F}H$ and $\mathbb{F}L$ are mutually inverse bijections. In order to simplify the notation, in what follows we will write:

$$\mathbb{F}L(t,v)=\big(t,\mathbb{F}L^{(2)}(t,v)\big),\quad \mathbb{F}H(t,p)=\big(t,\mathbb{F}H^{(2)}(t,p)\big),$$

so that $\mathbb{F}L^{(2)}$ and $\mathbb{F}H^{(2)}$ are respectively a TM^*-valued and a TM-valued map.

Let $L:U\subset I\!\!R\times TM\to I\!\!R$ be a Lagrangian on M and $\gamma:[a,b]\to M$ be a curve of class C^1, with $\big(t,\dot\gamma(t)\big)\in U$ for all t. The *action* $\mathcal{L}(\gamma)$ of L on the curve γ is given by the integral:

$$\mathcal{L}(\gamma)=\int_a^b L\big(t,\dot\gamma(t)\big)\,\mathrm{d}t. \tag{2.5}$$

$\mathcal{L}$ defines a functional on the set:

$$\begin{aligned}\Omega_{PQ}\big([a,b],M;U\big)&\\ &=\Big\{\gamma:[a,b]\xrightarrow{C^1}M:\gamma(a)\in P,\ \gamma(b)\in Q,\ \big(t,\dot\gamma(t)\big)\in U,\ \forall\,t\in[a,b]\Big\},\end{aligned} \tag{2.6}$$

where P and Q are two smooth embedded submanifolds of M. It is well known that $\Omega_{PQ}\big([a,b],M;U\big)$ has the structure of an infinite dimensional smooth Banach manifold (see for instance [15, 16]), and $\mathcal{L}$ is a functional of class C^1 on $\Omega_{PQ}\big([a,b],M;U\big)$. We will call $\mathcal{L}$ the *action functional* associated to the Lagrangian L.

We have the following characterization of the critical points of $\mathcal{L}$:

Proposition 2.7. *A curve $\gamma\in\Omega_{PQ}\big([a,b],M;U\big)$ is a critical point of $\mathcal{L}$ if and only if the following three conditions are satisfied:*

1. $\mathbb{F}L^{(2)}\big(a,\dot\gamma(a)\big)|_{T_{\gamma(a)}P}=0$ *and* $\mathbb{F}L^{(2)}\big(b,\dot\gamma(b)\big)|_{T_{\gamma(b)}Q}=0$*;*
2. $t\mapsto\mathbb{F}L\big(t,\dot\gamma(t)\big)$ *is of class* C^1*;*
3. *for all* $[t_0,t_1]\subset[a,b]$ *and for any chart* $q=(q_1,\dots,q_n)$ *on* M *whose domain contains* $\gamma\big([t_0,t_1]\big)$*, the* Euler–Lagrange equation *is satisfied in* $[t_0,t_1]$*:*

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial\dot q}\big(t,q(t),\dot q(t)\big)=\frac{\partial L}{\partial q}\big(t,q(t),\dot q(t)\big), \tag{2.7}$$

where $L(t,q,\dot q)$ denotes the coordinate representation of L.

Proof. Let $\gamma\in\Omega_{PQ}\big([a,b],M;U\big)$ be a critical point of $\mathcal{L}$. Let $[t_0,t_1]\subset[a,b]$ be an interval and consider a chart $q=(q_1,\dots,q_n)$ in M whose domain contains $\gamma\big([t_0,t_1]\big)$. Choose an arbitrary $v\in T_\gamma\Omega_{PQ}\big([a,b],M;U\big)$ with support contained in $]t_0,t_1[$; by standard computations it follows that:

$$\int_{t_0}^{t_1}\frac{\partial L}{\partial q}\big(t,q(t),\dot q(t)\big)\,v(t)+\frac{\partial L}{\partial\dot q}\big(t,q(t),\dot q(t)\big)\,\dot v(t)\,\mathrm{d}t=0. \tag{2.8}$$

[1] As a matter of fact, this same argument shows that $\mathbb{F}L:U\to V$ is a homeomorphism and therefore the Hamiltonian $H=L^*$ is continuous.

The fact that the equality above holds for every smooth v with support contained in $]t_0, t_1[$ implies that the term $\frac{\partial L}{\partial \dot q}\big(t, q(t), \dot q(t)\big)$ is of class C^1; this will follow[2] from the generalized functions calculus developed in Subsection 4.1 (see Corollary 4.5). Integration by parts in (2.8) and the Fundamental Lemma of Calculus of Variations imply then that equation (2.7) is satisfied. Observe also that the coordinate representation of the map $t \mapsto \mathbb{F}L^{(2)}\big(t, \dot\gamma(t)\big)$ is given by $t \mapsto \frac{\partial L}{\partial \dot q}\big(t, q(t), \dot q(t)\big)$, so that condition (2) is satisfied. Condition (1) follows easily by integrating by parts (2.8) in intervals of the form $[a, t_1]$ and $[t_0, b]$.

Conversely, if conditions (1), (2) and (3) are satisfied, equality (2.8) follows easily, which implies that $\mathrm{d}\mathcal{L}_\gamma(v) = 0$ for all $v \in T_\gamma\Omega_{PQ}\big([a,b], M; U\big)$ with small support. Since such v's span $T_\gamma\Omega_{PQ}\big([a,b], M; U\big)$, it follows that γ is a critical point of $\mathcal{L}$. $\square$

We now pass to the study of the Hamiltonian formalism, and we consider the *canonical symplectic form* ω on TM^*, given by $\omega = -\mathrm{d}\vartheta$, where the *canonical* 1*-form* ϑ on TM^* is defined by $\vartheta_p(\zeta) = p\big(\mathrm{d}\pi_p(\zeta)\big)$, for all $p \in TM^*$, $\zeta \in T_pTM^*$. If $q = (q_1, \ldots, q_n)$ is a chart on M and $(q, p) = (q_1, \ldots, q_n, p_1, \ldots, p_n)$ is the corresponding chart on TM^*, the forms ϑ and ω are given by:

$$\vartheta = \sum_{i=1}^n p_i \,\mathrm{d}q_i, \quad \omega = \sum_{i=1}^n \mathrm{d}q_i \wedge \mathrm{d}p_i. \tag{2.9}$$

Given a Hamiltonian H on M, we define its *Hamiltonian vector field* $\vec{H}$ to be the unique time-dependent vector field on TM^* satisfying:

$$\omega(\vec{H}, \cdot) = \mathrm{d}H_t,$$

where $H_t = H(t, \cdot)$.

We say that a curve $\gamma : [a, b] \to M$ is a *solution of the Hamiltonian* H if there exists a C^1-curve $\Gamma : [a, b] \to TM^*$ with $\pi \circ \Gamma = \gamma$ and such that:

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\Gamma(t) = \vec{H}\big(t, \Gamma(t)\big) \tag{2.10}$$

for all t. In this case, we say that Γ is a *Hamiltonian lift* of γ. In coordinates (q, p), equation (2.10) is written as:

$$\begin{cases} \dfrac{\mathrm{d}q}{\mathrm{d}t} = \dfrac{\partial H}{\partial p}\big(t, q(t), p(t)\big), \\[2ex] \dfrac{\mathrm{d}p}{\mathrm{d}t} = -\dfrac{\partial H}{\partial q}\big(t, q(t), p(t)\big). \end{cases} \tag{2.11}$$

These are called the *Hamilton equations* of H; observe that the first equation in (2.11) can be written intrinsically as:

$$\dot\gamma(t) = \mathbb{F}H^{(2)}\big(t, \Gamma(t)\big). \tag{2.12}$$

Theorem 2.8. *Let L be a hyper-regular Lagrangian on M and let $H = L^*$ be the corresponding hyper-regular Hamiltonian. Let P and Q be smooth submanifolds of M; a curve $\gamma \in \Omega_{PQ}\big([a, b], M; U\big)$ is a critical point of $\mathcal{L}$ if and only if γ is a solution of the Hamiltonian H which admits a Hamiltonian lift Γ such that*

$$\Gamma(a)|_{T_{\gamma(a)}P} = 0, \quad \Gamma(b)|_{T_{\gamma(b)}Q} = 0. \tag{2.13}$$

[2] Alternatively, one could use integration by parts and the fact that $\int_{t_0}^{t_1} \phi \dot v = 0$ for all smooth v with support in $]t_0, t_1[$ implies $\phi \equiv$ constant.

Proof. Let $\gamma \in \Omega_{PQ}\big([a,b],M;U\big)$ be a critical point of $\mathcal{L}$; set $\Gamma(t) = \mathbb{F}L^{(2)}\big(t,\dot\gamma(t)\big)$. Since $\mathbb{F}H$ and $\mathbb{F}L$ are mutually inverse, equation (2.12) follows. Moreover, by Proposition 2.7, Γ is of class C^1 and (2.13) holds. We now prove that the second Hamilton equation holds, using a chart (q,p) of TM^*. To this aim, we differentiate with respect to q the equality:

$$H\Big(t,q,\frac{\partial L}{\partial \dot q}(t,q,\dot q)\Big) = \frac{\partial L}{\partial \dot q}(t,q,\dot q)\,\dot q - L(t,q,\dot q),$$

obtaining:

$$\frac{\partial H}{\partial q}(t,q,p) + \frac{\partial H}{\partial p}(t,q,p)\,\frac{\partial^2 L}{\partial q\,\partial \dot q}(t,q,\dot q) = \frac{\partial^2 L}{\partial q\,\partial \dot q}(t,q,\dot q)\,\dot q - \frac{\partial L}{\partial q}(t,q,\dot q), \tag{2.14}$$

where $p = \frac{\partial L}{\partial \dot q}(t,q,\dot q)$. Using that $\mathbb{F}H$ and $\mathbb{F}L$ are mutually inverse, we get $\frac{\partial H}{\partial p}(t,q,p) = \dot q$; it follows from (2.14) that:

$$\frac{\partial H}{\partial q}(t,q,p) = -\frac{\partial L}{\partial q}(t,q,\dot q). \tag{2.15}$$

The second Hamilton equation now follows from formula (2.15) and from the Euler–Lagrange equation (2.7).

Conversely, suppose that γ is a solution of the Hamiltonian H which admits a Hamiltonian lift Γ satisfying (2.13). Since $\mathbb{F}H$ and $\mathbb{F}L$ are mutually inverse, from (2.12) it follows that $\Gamma(t) = \mathbb{F}L^{(2)}\big(t,\dot\gamma(t)\big)$. Finally, equality (2.15) and the second Hamilton equation imply the Euler–Lagrange equation (2.7), and the conclusion follows from Proposition 2.7. □

3. The Space of Horizontal Curves and its Differentiable Structure

In this section we recall some results concerning the manifold structure of the set of horizontal curves connecting two fixed submanifolds of a given manifold. Most of the material presented here is well-known in the context of sub-Riemannian geometry (see [3, 4, 10, 11, 12, 13]). Detailed proofs can be found in [17]. Actually, some minor adaptations of the proofs of [17] have to be made due to the fact that [17] deals with curves of Sobolev class H^1 while we have to deal here[3] with curves of class C^1.

Throughout the section we consider fixed an n-dimensional differentiable manifold M and a smooth distribution $\mathcal{D} \subset TM$ on M of rank $k \le n$. By a *horizontal curve* we mean a curve $\gamma : [a,b] \to M$ of class C^1 with $\gamma'(t) \in \mathcal{D}$ for all $t \in [a,b]$. Given smooth embedded submanifolds $P, Q \subset M$ we consider the following spaces:

$$\begin{aligned}
\Omega\big([a,b],M\big) &= \big\{\gamma : [a,b] \to M : \gamma \text{ is of class } C^1\big\};\\
\Omega_P\big([a,b],M\big) &= \big\{\gamma \in \Omega\big([a,b],M\big) : \gamma(a) \in P\big\};\\
\Omega_{PQ}\big([a,b],M\big) &= \big\{\gamma \in \Omega\big([a,b],M\big) : \gamma(a) \in P,\ \gamma(b) \in Q\big\};\\
\Omega\big([a,b],M,\mathcal{D}\big) &= \big\{\gamma \in \Omega\big([a,b],M\big) : \gamma \text{ is horizontal}\big\};\\
\Omega_P\big([a,b],M,\mathcal{D}\big) &= \Omega_P\big([a,b],M\big) \cap \Omega\big([a,b],M,\mathcal{D}\big);\\
\Omega_{PQ}\big([a,b],M,\mathcal{D}\big) &= \Omega_{PQ}\big([a,b],M\big) \cap \Omega\big([a,b],M,D\big).
\end{aligned}$$

[3]This is due to the fact that the *sub-Riemannian energy functional* studied in [17] is smooth on the space of H^1 curves while the action functional of an arbitrary Lagrangian is not in general even well-defined on such space.

It is well-known that $\Omega\big([a,b],M\big)$ has a natural structure of a Banach manifold (see for instance [15, 16]) and that $\Omega_P\big([a,b],M\big)$ and $\Omega_{PQ}\big([a,b],M\big)$ are embedded Banach submanifolds of $\Omega\big([a,b],M\big)$. Also $\Omega_P\big([a,b],M,\mathcal{D}\big)$ is an embedded Banach submanifold of $\Omega\big([a,b],M\big)$. The proof of this fact is obtained by using a suitable atlas for $\Omega\big([a,b],M\big)$ whose construction is described below.

If ξ is a vector bundle over M then a *time-dependent referential* of ξ over an open subset $A \subset I\!R \times M$ is a family $(X_i)_{i=1}^k$ of smooth maps $X_i : A \to \xi$ such that $\big(X_i(t,m)\big)_{i=1}^k$ is a basis of the fiber ξ_m for all $(t,m) \in A$. Given a time-dependent referential $(X_i)_{i=1}^n$ of the tangent bundle TM over an open subset $A \subset I\!R \times M$, we define a map:

$$\mathcal{B} : \Omega\big([a,b],M;\hat{A}\big) \longrightarrow C^0\big([a,b],I\!R^n\big),$$

by $\mathcal{B}(\gamma) = (h_1,\ldots,h_n)$, where:

$$\gamma'(t) = \sum_{i=1}^n h_i(t) X_i\big(t,\gamma(t)\big),$$

for all $t \in [a,b]$ and:

$$\hat{A} = \big\{(t,v) \in I\!R \times TM : \big(t,\pi(v)\big) \in A\big\}, \tag{3.1}$$

$$\Omega\big([a,b],M;\hat{A}\big) = \big\{\gamma \in \Omega\big([a,b],M\big) : \big(t,\gamma'(t)\big) \in \hat{A},\ \text{for all } t \in [a,b]\big\}. \tag{3.2}$$

Lemma 3.1. *If $\phi : U \subset M \to \widetilde{U} \subset I\!R^n$ is a local chart on M and $\mathcal{B}$ is defined as above then the map:*

(3.3)

$$\big\{\gamma \in \Omega\big([a,b],M;\hat{A}\big) : \gamma(a) \in U\big\} \ni \gamma \longmapsto \big(\phi(\gamma(a)),\mathcal{B}(\gamma)\big) \in I\!R^n \times C^0\big([a,b],I\!R^n\big),$$

is a local chart on the Banach manifold $\Omega\big([a,b],M\big)$.

Proof. It is a simple application of the Inverse Function Theorem on Banach manifolds (see [17, Corollary 4.2] for details on a similar construction). □

The proposition below implies that the local charts defined on Lemma 3.1 form an atlas for $\Omega\big([a,b],M\big)$.

Proposition 3.2. *Let ξ be a vector bundle over a differentiable manifold M. Given a continuous curve $\gamma : [a,b] \to M$, there exists a time-dependent referential $(X_i)_{i=1}^k$ of ξ whose domain A is an open neighborhood of the graph of γ in $I\!R \times M$, i.e., $\big(t,\gamma(t)\big) \in A$ for all $t \in [a,b]$.*

Proof. See [17, Lemma 2.3]. □

Using the atlas constructed above we can prove easily that $\Omega_P\big([a,b],M,\mathcal{D}\big)$ is a submanifold of $\Omega\big([a,b],M\big)$.

Proposition 3.3. *$\Omega_P\big([a,b],M,\mathcal{D}\big)$ is an embedded Banach submanifold of $\Omega\big([a,b],M\big)$.*

Proof. Applying Proposition 3.2 to the vector bundle $\mathcal{D}$ and to a complementary vector bundle of $\mathcal{D}$ in TM we obtain a time-dependent referential $(X_i)_{i=1}^n$ of TM such that $(X_i)_{i=1}^k$ is a time-dependent referential for $\mathcal{D}$; moreover, we may choose $(X_i)_{i=1}^n$ so that its domain $A \subset I\!R \times M$ contains the graph of any prescribed continuous curve in M. If ϕ is a local chart of M which sends P to an open subset of $I\!R^r \cong I\!R^r \times \{0\} \subset I\!R^n$ then the corresponding chart (3.3) on $\Omega\big([a,b],M\big)$ sends $\Omega_P\big([a,b],M,\mathcal{D}\big)$ to an open subset of $I\!R^r \times C^0\big([a,b],I\!R^k\big)$. □

Given Banach manifolds $\mathcal{M}, \mathcal{N}$, recall that a map $f : \mathcal{M} \to \mathcal{N}$ of class C^1 is said to be a *submersion* at a point $x \in \mathcal{M}$ if the differential $\mathrm{d}f_x : T_x\mathcal{M} \to T_{f(x)}\mathcal{N}$ is surjective and its Kernel $\mathrm{Ker}(\mathrm{d}f_x)$ is complemented in $T_x\mathcal{M}$, i.e., it admits a closed complementary subspace in $T_x\mathcal{M}$. When f is a submersion at x, then the intersection of $f^{-1}\big(f(x)\big)$ with some open neighborhood of x in $\mathcal{M}$ is a Banach submanifold of $\mathcal{M}$ whose tangent space at x is $\mathrm{Ker}(\mathrm{d}f_x)$. More generally, if $\mathcal{P} \subset \mathcal{N}$ is a Banach submanifold of $\mathcal{N}$ and $x \in f^{-1}(\mathcal{P})$ then we say that f is *transverse* to $\mathcal{P}$ at x if the composition of $\mathrm{d}f_x$ with the quotient map $T_{f(x)}\mathcal{N} \to T_{f(x)}\mathcal{N}/T_{f(x)}\mathcal{P}$ is surjective and has complemented kernel in $T_x\mathcal{M}$; equivalently, f is transverse to $\mathcal{P}$ at x if $\mathrm{Im}(\mathrm{d}f_x) + T_{f(x)}\mathcal{P} = T_{f(x)}\mathcal{N}$ and $\mathrm{d}f_x^{-1}(T_{f(x)}\mathcal{P})$ is complemented in $T_x\mathcal{M}$. If f is transverse to $\mathcal{P}$ at x then the intersection of $f^{-1}(\mathcal{P})$ with some open neighborhood of x in $\mathcal{M}$ is a Banach submanifold of $\mathcal{M}$ whose tangent space at x is $\mathrm{d}f_x^{-1}(T_{f(x)}\mathcal{P})$.

Definition 3.4. A curve $\gamma \in \Omega_{PQ}\big([a,b],M,\mathcal{D}\big)$ is called *regular* in $\Omega_{PQ}\big([a,b],M,\mathcal{D}\big)$ if the *endpoint map*:

$$\Omega_P\big([a,b],M,\mathcal{D}\big) \ni \mu \longmapsto \mu(b) \in M \tag{3.4}$$

is transverse to Q at the point γ. When γ is not regular in $\Omega_{PQ}\big([a,b],M,\mathcal{D}\big)$, we say that γ is *singular* in $\Omega_{PQ}\big([a,b],M,\mathcal{D}\big)$.

Since M is finite-dimensional, a curve γ is regular in $\Omega_{PQ}\big([a,b],M,\mathcal{D}\big)$ if and only if the image of the differential of (3.4) at γ plus $T_{\gamma(b)}Q$ equals $T_{\gamma(b)}M$.

Below we described an explicit method for computing the image of the differential of the endpoint map.

Definition 3.5. Denote by $\mathcal{D}^{\mathrm{o}} \subset TM^*$ the annihilator of $\mathcal{D}$. A curve $\eta : [a,b] \to TM^*$ of class C^1 is called a *characteristic* for $\mathcal{D}$ if $\eta\big([a,b]\big) \subset \mathcal{D}^{\mathrm{o}}$ and $\eta'(t) \in T_{\eta(t)}\mathcal{D}^{\mathrm{o}}$ belongs to the kernel of the restriction of $\omega_{\eta(t)}$ to $T_{\eta(t)}\mathcal{D}^{\mathrm{o}}$ (recall (2.9)).

Proposition 3.6. *The annihilator of the image of the differential of* (3.4) *at a curve γ is the subspace of $T_{\gamma(b)}M^*$ given by:*

$$\big\{\eta(b) : \eta \text{ is a characteristic of } \mathcal{D},\ \pi\circ\eta = \gamma,\ \eta(a)|_{T_{\gamma(a)}P} = 0\big\}.$$

Proof. The proof is a minor adaptation of the proof of [17, Theorem 4.9] where we consider the case that P is a point and we use H^1 curves instead of C^1 curves. □

Corollary 3.7. *A curve $\gamma \in \Omega_{PQ}\big([a,b],M,\mathcal{D}\big)$ is singular in $\Omega_{PQ}\big([a,b],M,\mathcal{D}\big)$ if and only if there exists a non zero characteristic $\eta : [a,b] \to TM^*$ of $\mathcal{D}$ with $\pi\circ\eta = \gamma$ and $\eta(a)|_{T_{\gamma(a)}P} = 0$, $\eta(b)|_{T_{\gamma(b)}Q} = 0$.*

Proof. Follows from Proposition 3.6 observing that a characteristic $\eta : [a,b] \to TM^*$ that vanishes at some $t_0 \in [a,b]$ is identically zero (see [17, Lemma 4.8]). □

Corollary 3.8. *If either $T_{\gamma(a)}P + \mathcal{D}_{\gamma(a)} = T_{\gamma(a)}M$ or $T_{\gamma(b)}Q + \mathcal{D}_{\gamma(b)} = T_{\gamma(b)}M$ then γ is regular in $\Omega_{PQ}\big([a,b],M,\mathcal{D}\big)$.* □

Remark 3.9. If the distribution $\mathcal{D}$ satisfies a strong non integrability condition (for instance, if $\mathcal{D}$ is a *contact distribution*) then the restriction of the symplectic form ω to the annihilator $\mathcal{D}^{\mathrm{o}}$ of $\mathcal{D}$ is nondegenerate outside the zero section and therefore all non zero characteristic curves of $\mathcal{D}$ are constant. In particular, every non constant curve in $\Omega_{PQ}\big([a,b],M,\mathcal{D}\big)$ is regular.

So far we have looked at $\Omega_{PQ}\big([a,b],M,\mathcal{D}\big)$ as the set of curves γ in the Banach manifold $\Omega_P\big([a,b],M,\mathcal{D}\big)$ satisfying the constraint $\gamma(b)\in Q$. We could also think of $\Omega_{PQ}\big([a,b],M,\mathcal{D}\big)$ as the set of curves in the Banach manifold $\Omega_{PQ}\big([a,b],M\big)$ satisfying the constraint $\mathrm{Im}(\gamma')\subset\mathcal{D}$. Actually, the latter point of view will be needed in the proof of our main theorem in Subsection 4.2. Our goal now is to show that both constraints have the same singularities. This fact was shown in [17] in the context of curves of class H^1. However, in the case of curves of class C^1 the problem is a little harder due to the fact that not every closed subspace of a Banach space is complemented. We have thus decided to give all the details of the proof.

The lemma below is a general principle that says that if a set is defined by two constraints then the singularities of the first in the space defined by the second constraint equals the singularities of the second in the space defined by the first.

Lemma 3.10. *Let $\mathcal{M}$, $\mathcal{N}_1$, $\mathcal{N}_2$ be Banach manifolds and $\mathcal{P}_1\subset\mathcal{N}_1$, $\mathcal{P}_2\subset\mathcal{N}_2$ be Banach submanifolds. Assume that we are given maps $f_i:\mathcal{M}\to\mathcal{N}_i$, $i=1,2$, of class C^1 and a point $x\in f_1^{-1}(\mathcal{P}_1)\cap f_2^{-1}(\mathcal{P}_2)$ such that f_i is transverse to $\mathcal{P}_i$ at x, $i=1,2$. Then the restriction $f_1|_{f_2^{-1}(\mathcal{P}_2)}$ is transverse to $\mathcal{P}_1$ at x if and only if the restriction $f_2|_{f_1^{-1}(\mathcal{P}_1)}$ is transverse to $\mathcal{P}_2$ at x.*

Proof. Consider the Banach spaces $X=T_x\mathcal{M}$, $Y_i=T_{f_i(x)}\mathcal{N}_i/T_{f_i(x)}\mathcal{P}_i$, $i=1,2$, and the continuous linear maps $L_i:X\to Y_i$, $i=1,2$, given by composition of $\mathrm{d}f_i(x)$ with the quotient map $T_{f_i(x)}\mathcal{N}_i\to T_{f_i(x)}\mathcal{N}_i/T_{f_i(x)}\mathcal{P}_i$. We know that both L_1 and L_2 are surjective and have complemented kernel. We have to show that $L_1|_{\mathrm{Ker}(L_2)}$ is surjective with complemented kernel if and only if $L_2|_{\mathrm{Ker}(L_1)}$ is surjective with complemented kernel. To this aim, observe first that $L_1|_{\mathrm{Ker}(L_2)}$ is surjective if and only if $\mathrm{Ker}(L_1)+\mathrm{Ker}(L_2)=X$ and the latter condition is symmetric in L_1 and L_2. Finally, to complete the proof we show that, given $i=1,2$, then $\mathrm{Ker}(L_1)\cap\mathrm{Ker}(L_2)$ is complemented in $\mathrm{Ker}(L_i)$ if and only if it is complemented in X. If $\mathrm{Ker}(L_1)\cap\mathrm{Ker}(L_2)$ is complemented in X then by intersecting a closed complement of $\mathrm{Ker}(L_1)\cap\mathrm{Ker}(L_2)$ in X with $\mathrm{Ker}(L_i)$ we obtain a closed complement of $\mathrm{Ker}(L_1)\cap\mathrm{Ker}(L_2)$ in $\mathrm{Ker}(L_i)$. Conversely, if Z is a closed complement of $\mathrm{Ker}(L_1)\cap\mathrm{Ker}(L_2)$ in $\mathrm{Ker}(L_i)$ and Z' is a closed complement of $\mathrm{Ker}(L_i)$ in X then $Z\oplus Z'$ is a closed complement of $\mathrm{Ker}(L_1)\cap\mathrm{Ker}(L_2)$ in X because $X=\mathrm{Ker}(L_i)\oplus Z'$ has the product topology of $\mathrm{Ker}(L_i)$ and Z'. □

We can now prove the following:

Proposition 3.11. *Let $(\theta_i)_{i=1}^{n-k}$ be a time-dependent referential of $\mathcal{D}^{\mathrm{o}}$ defined over an open subset $A\subset I\!R\times M$; set $\theta=(\theta_1,\ldots,\theta_{n-k})$, so that $\theta_{(t,m)}:T_mM\to I\!R^{n-k}$ is a surjective linear map with $\mathrm{Ker}(\theta_{(t,m)})=\mathcal{D}_m$ for all $(t,m)\in A$. Consider the map:*

$$\Theta:\Omega_{PQ}\big([a,b],M;\hat{A}\big)\longrightarrow C^0\big([a,b],I\!R^{n-k}\big)$$

defined by:

$$\Theta(\gamma)(t)=\theta\big(\gamma'(t)\big),\quad t\in[a,b].$$

Then γ is regular in $\Omega_{PQ}\big([a,b],M,\mathcal{D}\big)$ (in the sense of Definition 3.4) if and only if Θ is a submersion at γ.

Proof. Let $\overline{\Theta}$ denote the extension of Θ to

$$\Omega_P\big([a,b],M;\hat{A}\big)=\Omega_P\big([a,b],M\big)\cap\Omega\big([a,b],M;\hat{A}\big)$$

which is again defined by $\overline{\Theta}(\gamma)(t) = \theta\big(\gamma'(t)\big)$. The conclusion will follow by applying Lemma 3.10 with $\mathcal{M} = \Omega_P\big([a,b], M; \hat{A}\big)$, $\mathcal{N}_1 = C^0\big([a,b], I\!R^{n-k}\big)$, $\mathcal{P}_1 = \{0\}$, $\mathcal{N}_2 = M$, $\mathcal{P}_2 = Q$, $f_1 = \overline{\Theta}$ and $f_2 : \mathcal{M} \to \mathcal{N}_2$ equal to the endpoint map $\mu \mapsto \mu(b)$. Since f_2 is obviously a submersion, we only need to show that $f_1 = \overline{\Theta}$ is a submersion. Choose a distribution $\mathcal{D}' \subset TM$ with $TM = \mathcal{D} \oplus \mathcal{D}'$ and let $(X_i)_{i=1}^{n-k}$ be the time-dependent referential of $\mathcal{D}'$ over A which is dual to $(\theta_i)_{i=1}^{n-k}$, i.e., $\theta_i(X_j) = 1$ for $i = j$ and $\theta_i(X_j) = 0$ for $i \neq j$. Choose a time-dependent referential $(X_i)_{i=n-k+1}^{n}$ of $\mathcal{D}$ over an open neighborhood of the graph of γ. The coordinate representation of Θ in the chart (3.3) corresponding to $(X_i)_{i=1}^{n}$ is the natural projection of $I\!R^n \oplus C^0\big([a,b], I\!R^n\big)$ onto $C^0\big([a,b], I\!R^{n-k}\big)$. This shows that $\overline{\Theta}$ is a submersion and concludes the proof. □

4. Lagrangians with linear constraints and degenerate Hamiltonians

Let M be an n-dimensional manifold and $\mathcal{D} \subset TM$ be a smooth distribution of rank k. We consider $\mathcal{D}$ as a vector bundle over M with projection $\pi : \mathcal{D} \to M$. We apply the theory of Subsection 2.1 to the vector bundle $\xi = I\!R \times \mathcal{D}$ over the manifold $I\!R \times M$, with projection $\mathrm{Id} \times \pi$. The fiber $\xi_{(t,m)}$ is given by $\{t\} \times \mathcal{D}_m$.

Let $L : U \subset \xi \to I\!R$ be a map of class C^2 defined in an open set $U \subset \xi$; we assume that L is hyper-regular in the sense of Definition 2.4, so that (by the Inverse Function Theorem) the fiber derivative $\mathbb{F}L : U \to V$ is a C^1 diffeomorphism onto an open subset $V \subset \xi^*$. Let $H_0 = L^*$ be the Legendre transform of L. Then $H_0 : V \to I\!R$ is a map of class C^1 whose restriction to each fiber of ξ^* is of class C^2; moreover, the fiber derivative $\mathbb{F}H_0 : V \to U$ is the inverse of $\mathbb{F}L$ (see Proposition 2.5).

For every $p \in TM^*$ we denote by $p|_{\mathcal{D}}$ the restriction of $p \in T_mM^*$ to $\mathcal{D}_m$. Observe that the *restriction map* $TM^* \ni p \to p|_{\mathcal{D}} \in \mathcal{D}^*$ is the transpose of the vector bundle inclusion $\mathcal{D} \to TM$. By composing H_0 with the restriction map $TM^* \to \mathcal{D}^*$ we obtain a map $H : \widetilde{V} \to I\!R$ given by:

$$H(t,p) = H_0\big(t, p|_{\mathcal{D}}\big), \quad (t,p) \in \widetilde{V}, \tag{4.1}$$

where:

$$\widetilde{V} = \big\{(t,p) \in I\!R \times TM^* : (t, p|_{\mathcal{D}}) \in V\big\}.$$

Observe that H is a Hamiltonian on M (see Definition 2.6) of class C^1 defined in the open set $\widetilde{V} \subset I\!R \times TM^*$. We will call L a *constrained Lagrangian* on M, and H the corresponding *degenerate Hamiltonian* (observe indeed that H cannot be regular unless $\mathcal{D} = TM$).

Given any two submanifolds P and Q of M then a constrained Lagrangian L on M defines an action functional $\mathcal{L}$ on $\Omega_{PQ}\big([a,b], M, \mathcal{D}; U\big)$ by formula (2.5). Our goal is to determine the critical points of $\mathcal{L}$.

The following is the main result of the paper and its proof is given in Subsection 4.2:

Theorem 4.1. *Let M be an n-dimensional manifold, $\mathcal{D} \subset TM$ be a smooth distribution of rank k and $L : U \subset I\!R \times \mathcal{D} \to I\!R$ be a hyper-regular constrained Lagrangian of class C^2. Let $H_0 = L^*$ be the Legendre transform of L and let H be the corresponding degenerate Hamiltonian as in* (4.1)*.*

Fix two submanifolds P and Q of M and let $\mathcal{L}$ be the action functional of L defined in the space $\Omega_{PQ}\big([a,b], M, \mathcal{D}; U\big) = \Omega_{PQ}\big([a,b], M, \mathcal{D}\big) \cap \Omega_{PQ}\big([a,b], M; U\big)$, given by (2.5)*. Let $\gamma \in \Omega_{PQ}\big([a,b], M, \mathcal{D}; U\big)$ be a regular curve. Then, γ is a critical point of $\mathcal{L}$ if and only if it is a solution of H that admits a Hamiltonian lift $\Gamma : [a,b] \to TM^*$ with $\Gamma(a)|_{T_{\gamma(a)}P} = 0$ and $\Gamma(b)|_{T_{\gamma(b)}Q} = 0$.*

The classical example of a constrained hyper-regular Lagrangian function L is given by:

$$L(t,v)=\tfrac{1}{2}\,g(v,v)-V\big(\pi(v)\big), \tag{4.2}$$

where g is a ***sub-Riemannian metric*** on $(M,\mathcal{D})$ (i.e., a smooth Riemannian structure on the vector bundle $\mathcal{D}$) and $V:M\to I\!R$ is a map of class C^2. The fiber derivative $\mathbb{F}L$ of (4.2) is given by:

$$\mathbb{F}L(t,v)=g(v,\cdot)\in\mathcal{D}^*,$$

so that L is indeed hyper-regular. Recalling (2.4), we compute as follows:

$$\begin{aligned}E_L(t,v)&=\tfrac{1}{2}\,g(v,v)+V\big(\pi(v)\big),\quad v\in\mathcal{D},\\ H_0(t,\rho)&=\tfrac{1}{2}\,g^{-1}(\rho,\rho)+V\big(\pi(\rho)\big),\quad \rho\in\mathcal{D}^*,\end{aligned}$$

where g^{-1} denotes the induced Riemannian structure on the dual bundle $\mathcal{D}^*$. The degenerate Hamiltonian H corresponding to (4.2) is thus given by:

$$H(t,p)=\tfrac{1}{2}\,g^{-1}(p|_{\mathcal{D}},p|_{\mathcal{D}})+V\big(\pi(p)\big),\quad p\in TM^*.$$

Theorem 4.1 implies that the critical points of the action functional $\mathcal{L}$ corresponding to (4.2) on the space $\Omega_{PQ}\big([a,b],M,\mathcal{D}\big)$ are the solutions of H that admit a Hamiltonian lift $\Gamma:[a,b]\to TM^*$ satisfying the boundary conditions $\Gamma(a)|_{T_{\gamma(a)}P}=0$ and $\Gamma(b)|_{T_{\gamma(b)}Q}=0$. Observe that (in the case when P and Q are points) we obtain the equations for the trajectories of the Vakonomic mechanics given in [8]; when $V=0$ we obtain the equations for the normal geodesics of the sub-Riemannian manifold $(M,\mathcal{D},g)$ (see [10]).

Remark 4.2. We emphasize that, in general, a minimum of the action functional $\mathcal{L}$ may not be a regular curve in $\Omega_{PQ}\big([a,b],M,\mathcal{D}\big)$, and in this situation it may not satisfy the Hamilton equations of H. Examples of this phenomenon are given in [10, 11] in the sub-Riemannian case $L(t,v)=\frac{1}{2}g(v,v)$. Hence, one can only conclude that a minimum of $\mathcal{L}$ is either a solution of the Hamilton equations or the projection of a non null characteristic of $\mathcal{D}$.

4.1. *Generalized functions calculus*

For the proof of Theorem 4.1 we will occasionally have to consider derivatives of functions that are in principle only continuous[4]. These derivatives should be understood in the sense of Schwartz's generalized functions calculus. However, the usual definition of the generalized functions space as the dual of the space of smooth compactly supported maps only allows products of generalized functions by smooth maps. To overcome this difficulty, we introduce a calculus for generalized functions of *stronger* regularity, that are elements of the dual of a space of functions with *weaker* regularity.

Let V be a real finite dimensional vector space. For $k\ge 0$, we define $C^k_0\big([a,b],V\big)$ to be the Banach space of V-valued C^k maps on $[a,b]$ whose first k derivatives vanish at a and at b; we endow it with the standard C^k-norm. We denote by $D^k\big([a,b],V\big)$ the dual Banach space of $C^k_0\big([a,b],V^*\big)$ (dual spaces will *always* be meant in the *topological sense*). Denoting by $L^p\big([a,b],V\big)$ the Banach space of V-valued measurable functions on $[a,b]$ whose p-th power is Lebesgue integrable, we have an inclusion:

$$L^1\big([a,b],V\big)\hookrightarrow D^k\big([a,b],V\big) \tag{4.3}$$

[4]This situation already occurred in the proof of Proposition 2.7. In that case the difficulty could also be circumvented by a simpler technique.

defined by

$$\langle f,\alpha\rangle=\int_a^b \alpha(t)\,f(t)\,\mathrm{d}t,\quad f\in L^1\big([a,b],V\big),\ \alpha\in C_0^k\big([a,b],V^*\big);$$

in the formula above we have denoted by $\langle f,\alpha\rangle$ the evaluation at α of the linear functional which is the image of f by (4.3). In what follows we will always identify a function $f\in L^1\big([a,b],V\big)$ with its image by (4.3); moreover, the evaluation of $f\in D^k\big([a,b],V\big)$ at $\alpha\in C_0^k\big([a,b],V^*\big)$ will always be denoted by $\langle f,\alpha\rangle$.

Observe that we have inclusions $D^k\hookrightarrow D^{k+1}$ defined by restriction of the functionals, i.e., $D^k\hookrightarrow D^{k+1}$ is simply the transpose of the inclusion of $C_0^{k+1}\big([a,b],V^*\big)$ in $C_0^k\big([a,b],V^*\big)$.

We summarize the observations above by the following diagram:

$$\cdots\hookrightarrow C^1\hookrightarrow C^0\hookrightarrow L^1\hookrightarrow D^0\hookrightarrow D^1\hookrightarrow\cdots$$

An element f of any space $D^k\big([a,b],V\big)$ is called a *generalized function*. In what follows, we will occasionally write simply C^k, C_0^k, D^k, L^p instead of $C^k\big([a,b],V\big)$, $C_0^k\big([a,b],V\big)$, $D^k\big([a,b],V\big)$, $L^p\big([a,b],V\big)$.

In addition to the standard vector space operations in D^k, we define the following:

- *derivative operation:* for $f\in D^k\big([a,b],V\big)$, we define the *derivative* of f to be the generalized function $f'\in D^{k+1}\big([a,b],V\big)$ defined by:

  $$\langle f',\alpha\rangle=-\langle f,\alpha'\rangle,$$

  for all $\alpha\in C_0^{k+1}\big([a,b],V^*\big)$;
- *product operation:* for $f\in D^k\big([a,b],V\big)$, $g\in C^k\big([a,b],W\big)$ and a fixed bilinear map $V\times W\to U$, we define the product $fg\in D^k\big([a,b],U\big)$ as follows. The bilinear map $V\times W\to U$ induces a bilinear map $W\times U^*\to V^*$ defined by $(w\cdot u^*)(v)=u^*(v\cdot w)$; we set:

  $$\langle fg,\alpha\rangle=\langle f,g\cdot\alpha\rangle,$$

  for all $\alpha\in C_0^k\big([a,b],U^*\big)$;
- *restriction operation:* for $f\in D^k\big([a,b],V\big)$ and $[c,d]\subset[a,b]$, we set:

  $$\langle f\vert_{[c,d]},\alpha\rangle=\langle f,\overline{\alpha}\rangle,$$

  for all $\alpha\in C_0^k\big([c,d],V^*\big)$, where $\overline{\alpha}\in C_0^k\big([a,b],V^*\big)$ is the extension to zero of α outside $[c,d]$.

It is easily seen that when we apply the above operations to elements of D^k which correspond to functions then we obtain the standard operations on functions. Moreover, the standard Leibnitz rule for derivatives of products holds for generalized functions, i.e.:

$$(fg)'=f'g+fg',$$

for all $f\in D^k$ and $g\in C^{k+1}$.

In order to prove some regularity results we present the following elementary lemmas.

Lemma 4.3. *Let $f\in D^k\big([a,b],V\big)$ be such that $f'=0$. Then f is a constant function.*

Proof. We first consider the case $V=I\!R$. If $f'=0$, then $\langle f,\alpha'\rangle=0$ for all $\alpha\in C_0^{k+1}\big([a,b],I\!R\big)$, hence $\langle f,\beta\rangle=0$ for all $\beta\in C_0^k\big([a,b],I\!R\big)$ with $\int_a^b\beta=0$. Choose $\beta_0\in C_0^k\big([a,b],I\!R\big)$ with $\int_a^b\beta_0=1$; set $c=\langle f,\beta_0\rangle$. It is easily seen that $f\equiv c$.

For the general case, observe that for all $\lambda \in V^*$, the product $\lambda f \in D^k\big([a,b],\mathbb{R}\big)$ has vanishing derivative, and hence it is constant. Since λ is arbitrary, it follows that f is constant. $\square$

Lemma 4.4. *Let $f \in D^k\big([a,b],V\big)$, $k \geq 1$; there exists an element $F \in D^{k-1}\big([a,b],V\big)$ with $F' = f$. If $f \in D^0\big([a,b],V\big)$, there exists $F \in L^2\big([a,b],V\big)$ with $F' = f$.*

Proof. Consider the map $\mathrm{d} : C_0^{k+1} \to C_0^k$ given by $\mathrm{d}(\alpha) = \alpha'$. It is easily seen that d is injective with closed image. It follows that the *transpose map* $\mathrm{d}^* : D^k \to D^{k+1}$ is surjective; clearly, the derivative operator for generalized functions is $-\mathrm{d}^*$, which proves the first part of the thesis.

For the case $k = 0$, let H_0^1 denote the Sobolev space of absolutely continuous functions $\alpha : [a,b] \to V^*$ having square integrable derivative, and such that $\alpha(a) = \alpha(b) = 0$. Again, the derivation map $\mathrm{d} : H_0^1 \to L^2$ is injective and has closed image. Therefore, given $f \in D^0$, we can find $F \in L^{2^*} \simeq L^2$ with $\mathrm{d}^* F = -f|_{H_0^1}$. It follows that $F' = f$. $\square$

Corollary 4.5 (Bootstrap lemma)**.** *Let f be a generalized function.*

(1) *If $f' \in D^0$ then $f \in L^2$;*
(2) *If $f' \in L^2$ then $f \in C^0$;*
(3) *If $f' \in C^0$ then $f \in C^1$.*

Proof. We prove, for example, the first item. By Lemma 4.4, we can find $F \in L^2$ with $F' = f'$. By Lemma 4.3, it follows that $F - f$ is constant, hence $f \in L^2$.

The other items are proven similarly. $\square$

We now give a result that shows that *regularity* of a generalized function is a local property:

Lemma 4.6. *Let λ be a generalized function on $[a,b]$. Suppose that for all $t \in [a,b]$ there exists $\varepsilon > 0$ such that the restriction $\lambda|_{[t-\varepsilon,t+\varepsilon]\cap[a,b]}$ is of class C^k, $k \geq 0$. Then λ is of class C^k.*

Proof. Consider a partition $a = t_0 < t_1 < \ldots < t_r = b$ such that $f_i = \lambda|_{[t_i,t_{i+2}]}$ is of class C^k for all $i = 0,\ldots,r-2$. Since the operation of restriction for generalized functions gives the standard operation of restriction for functions, it follows that:

$$f_i|_{[t_{i+1},t_{i+2}]} = \lambda|_{[t_{i+1},t_{i+2}]} = f_{i+1}|_{[t_{i+1},t_{i+2}]},$$

for $i = 0,\ldots,r-3$. Hence there exists a C^k map f on $[a,b]$ such that $f|_{[t_i,t_{i+2}]} = f_i$ for all $i = 0,\ldots,r-2$. We know that $\langle f,\alpha\rangle = \langle\lambda,\alpha\rangle$ if α has support contained in some interval $]t_i,t_{i+2}[$; but such α's span a dense subspace of the domain of the linear functional λ and therefore $\lambda = f$. $\square$

Finally, we need the following result that relates the dual spaces of C^0 and C^0_0. For $t \in [a,b]$ and $\sigma \in V$, we denote by $\delta_t^\sigma \in C^0\big([a,b],V^*\big)^*$ the *Dirac's delta*, defined by:

$$\langle\delta_t^\sigma,\alpha\rangle = \alpha(t)\,\sigma, \quad \alpha \in C^0\big([a,b],V^*\big).$$

Lemma 4.7. *If $\lambda \in C^0\big([a,b],V^*\big)^*$ vanishes identically on $C_0^0\big([a,b],V^*\big)$ then there exist σ_a and σ_b in V such that:*

$$\lambda = \delta_a^{\sigma_a} + \delta_b^{\sigma_b}. \tag{4.4}$$

Proof. If $\mathcal{A}$ denotes the subspace of $C^0\big([a,b],V^*\big)$ consisting of *affine maps* $\alpha(t)=Pt+Q$ then obviously:

$$C^0\big([a,b],V^*\big)=C^0_0\big([a,b],V^*\big)\oplus\mathcal{A}.$$

It is easy to see that we can find $\sigma_a,\sigma_b\in V$ such that both sides of (4.4) agree on $\mathcal{A}$. Since both sides of (4.4) vanish on $C^0_0\big([a,b],V^*\big)$, the conclusion follows. □

4.2. *Proof of Theorem 4.1*

The proof of Theorem 4.1 is based on the method of Lagrange multipliers, and we start with the precise statement of the result needed for our purposes.

Proposition 4.8. *Let $\mathcal{M}$ be a Banach manifold, E a Banach space and $F:\mathcal{M}\to I\!R$, $g:\mathcal{M}\to E$ maps of class C^1. Let $p\in g^{-1}(0)$ be such that g is a submersion at p. Then, p is a critical point for $f|_{g^{-1}(0)}$ if and only if there exists $\lambda\in E^*$ such that p is a critical point for the functional $f_\lambda=f-\lambda\circ g$ in $\mathcal{M}$.*

Proof. The point p is critical for $f|_{g^{-1}(0)}$ if and only if $\mathrm{d}f(p)$ vanishes on $T_pg^{-1}(0)=\mathrm{Ker}\big(\mathrm{d}g(p)\big)$. The conclusion follows from elementary functional analysis arguments. □

The linear functional $\lambda\in E^*$ of Proposition 4.8 is called the *Lagrange multiplier* of the constrained critical point p; it is easily seen that such λ is unique. We can now prove the main result of the section. In the argument we will need a regularity result for a Lagrangian multiplier; such proof is postponed to Lemma 4.9.

Proof of Theorem 4.1. We start by choosing an arbitrary complementary distribution $\mathcal{D}'$ to $\mathcal{D}$, i.e., a smooth distribution of rank $n-k$ in M such that $T_mM=\mathcal{D}_m\oplus\mathcal{D}'_m$ for all $m\in M$; moreover, we fix an arbitrary smooth Riemannian structure g on the vector bundle $\mathcal{D}'$. Let $\pi_{\mathcal{D}}:TM\to\mathcal{D}$ and $\pi_{\mathcal{D}'}:TM\to\mathcal{D}'$ denote the projections and define an extension $\widetilde{L}:\widetilde{U}\subset I\!R\times M\to I\!R$ of L by:

$$\widetilde{L}(t,v)=L\big(t,\pi_{\mathcal{D}}(v)\big)+\tfrac{1}{2}\,g\big(\pi_{\mathcal{D}'}(v),\pi_{\mathcal{D}'}(v)\big),\tag{4.5}$$

where

$$\widetilde{U}=\Big\{(t,v)\in I\!R\times TM:\big(t,\pi_{\mathcal{D}}(v)\big)\in U\Big\}.$$

Then $\widetilde{U}$ is open in $I\!R\times TM$ and $\widetilde{L}$ is a Lagrangian on M as in Definition 2.6; we denote by $\widetilde{\mathcal{L}}$ the corresponding action functional in $\Omega_{PQ}\big([a,b],M;\widetilde{U}\big)$, defined as in (2.5).

Let θ, Θ, A and $\hat{A}$ be as in the statement of Proposition 3.11 (recall also (3.1) and (3.2)). Then, since γ is regular, the map Θ is a submersion at γ; moreover, γ is a critical point of $\mathcal{L}$ in $\Omega_{PQ}\big([a,b],M,\mathcal{D};U\big)$ if and only if it is a critical point of $\widetilde{\mathcal{L}}|_{\Theta^{-1}(0)}$. By the method of Lagrange multipliers (Proposition 4.8), this is equivalent to the existence of $\lambda\in C^0\big([a,b],I\!R^{n-k}\big)^*$ such that γ is a critical point of $\widetilde{\mathcal{L}}_\lambda=\widetilde{\mathcal{L}}-\lambda\circ\Theta$ in $\Omega_{PQ}\big([a,b],M;\hat{A}\cap\widetilde{U}\big)$.

We will prove in Lemma 4.9 below that the Lagrange multiplier λ is of class C^1, i.e., that it is given by:

$$\lambda(\alpha)=\int_a^b\lambda_0(t)\,\alpha(t)\,\mathrm{d}t,\quad\forall\,\alpha\in C^0\big([a,b],I\!R^{n-k}\big),\tag{4.6}$$

for some C^1 map $\lambda_0 : [a,b] \to (I\!R^{n-k})^*$. Therefore, $\widetilde{\mathcal{L}}_\lambda$ is the action functional corresponding to the Lagrangian $\widetilde{L}_\lambda$ in M defined by:

$$\widetilde{L}_\lambda(t,v) = \widetilde{L}(t,v) - \lambda_0(t)\,\theta_{(t,m)}(v), \quad (t,v) \in \hat{A} \cap \widetilde{U}, \tag{4.7}$$

where $m = \pi(v)$.

We now prove that $\widetilde{L}$ and $\widetilde{L}_\lambda$ are hyper-regular and we compute their Legendre transforms. The fiber derivatives $\mathbb{F}\widetilde{L}$ and $\mathbb{F}\widetilde{L}_\lambda$ are easily computed as:

$$\mathbb{F}\widetilde{L}(t,v) = \mathbb{F}L\big(t, \pi_{\mathcal{D}}(v)\big) \circ \pi_{\mathcal{D}} + g\big(\pi_{\mathcal{D}'}(v), \pi_{\mathcal{D}'}(\,\cdot\,)\big) \in T_mM^*, \tag{4.8}$$

$$\mathbb{F}\widetilde{L}_\lambda(t,v) = \mathbb{F}\widetilde{L}(t,v) - \lambda_0(t)\,\theta_{(t,m)} \in T_mM^*, \tag{4.9}$$

where $m = \pi(v)$. The hyper-regularity is proven by exhibiting explicit inverses:

$$\begin{aligned} \mathbb{F}\widetilde{L}^{-1}(t,p) &= \mathbb{F}L^{-1}(t, p|_{\mathcal{D}}) + g^{-1}(p|_{\mathcal{D}'}), \\ \mathbb{F}\widetilde{L}_\lambda^{-1}(t,p) &= \mathbb{F}\widetilde{L}^{-1}\big(t, p + \lambda_0(t)\,\theta_{(t,m)}\big); \end{aligned} \tag{4.10}$$

by g^{-1} in the above formula we mean the inverse of g seen as a linear map from $\mathcal{D}_m$ to $\mathcal{D}_m^*$.

We now compute the Legendre transforms $\widetilde{H}$ and $\widetilde{H}_\lambda$ of $\widetilde{L}$ and $\widetilde{L}_\lambda$ respectively. Using Definition 2.1 and equations (4.8), (4.9), we compute easily:

$$E_{\widetilde{L}_\lambda}(t,v) = E_{\widetilde{L}}(t,v) = E_L\big(t, \pi_{\mathcal{D}}(v)\big) + \tfrac{1}{2}\, g\big(\pi_{\mathcal{D}'}(v), \pi_{\mathcal{D}'}(v)\big); \tag{4.11}$$

and, using (4.10), we therefore obtain:

$$\begin{aligned} \widetilde{H}(t,p) &= H(t,p) + \tfrac{1}{2}\, g^{-1}(p|_{\mathcal{D}'}, p|_{\mathcal{D}'}), \\ \widetilde{H}_\lambda(t,p) &= \widetilde{H}(t, p + \lambda_0(t)\,\theta_{(t,m)}) \\ &= H(t,p) + \tfrac{1}{2}\, g^{-1}\big((p + \lambda_0(t)\,\theta_{(t,m)})|_{\mathcal{D}'}, (p + \lambda_0(t)\,\theta_{(t,m)})|_{\mathcal{D}'}\big). \end{aligned}$$

We now compute the Hamilton equations of the Hamiltonian $\widetilde{H}_\lambda$ with the help of local coordinates $(q_1, \ldots, q_n, p_1, \ldots, p_n)$ in TM^* and of a local g-orthonormal referential $X_1, \ldots, X_{n-k}$ of $\mathcal{D}'$.

We write:

$$\widetilde{H}_\lambda(t,p) = H(t,p) + \frac{1}{2} \sum_{i=1}^{n-k} \big(p + \lambda_0(t)\,\theta_{(t,m)}\big)(X_i)^2, \tag{4.12}$$

and, using (2.11), the Hamilton equations of $\widetilde{H}_\lambda$ are given by:

$$\begin{cases} \dfrac{\mathrm{d}q}{\mathrm{d}t} = \dfrac{\partial H}{\partial p} + \displaystyle\sum_{i=1}^{n-k} (p + \lambda_0\,\theta)(X_i)\,X_i, \\[2ex] \dfrac{\mathrm{d}p}{\mathrm{d}t} = -\dfrac{\partial H}{\partial q} - \displaystyle\sum_{i=1}^{n-k} (p + \lambda_0\,\theta)(X_i)\,\left[\lambda_0\,\dfrac{\partial \theta}{\partial q}(X_i) + (p + \lambda_0\,\theta)\left(\dfrac{\partial X_i}{\partial q}\right)\right]. \end{cases} \tag{4.13}$$

By Theorem 2.8, γ is a critical point of $\widetilde{\mathcal{L}}_\lambda$ if and only if it admits a lift $\Gamma : [a,b] \to TM^*$ satisfying (4.13) with $\Gamma(a) \in T_{\gamma(a)}P^o$ and $\Gamma(b) \in T_{\gamma(b)}Q^o$.

Now, it follows easily from (4.1) that $\frac{\partial H}{\partial p}$ is in $\mathcal{D}$; since γ is horizontal, i.e., $\frac{\mathrm{d}q}{\mathrm{d}t} \in \mathcal{D}$, from the first equation of (4.13) it follows that $(p + \lambda_0\,\theta)(X_i) = 0$ for all $i = 1, \ldots, n-k$. Setting $(p + \lambda_0\,\theta)(X_i) = 0$ in (4.13) we obtain the Hamilton equations of H, which concludes the proof. $\square$

We are left with the proof of the *regularity* of the Lagrange multiplier λ. We will use the generalized functions calculus developed in Subsection 4.1.

Lemma 4.9. *Under the assumptions of Theorem 4.1, using the notations adopted in its proof, if γ is horizontal and if, for some $\lambda \in C^0\big([a,b],\mathbb{R}^{n-k}\big)^*$, it is a critical point of $\widetilde{\mathcal{L}} - \lambda\circ\Theta$, then there exists a C^1 map $\lambda_0 : [a,b] \to (\mathbb{R}^{n-k})^*$ such that* (4.6) *holds.*

Proof. We set

$$\lambda_0 = \lambda|_{C^0_0([a,b],\mathbb{R}^{n-k})} \in D^0\big([a,b],(\mathbb{R}^{n-k})^*\big);$$

we first prove the regularity of the generalized function λ_0. To this aim, we *localize* the problem by considering variational vector fields along γ having support in the domain of a local chart $q = (q_1,\ldots,q_n)$ in M.

Let $[c,d] \subset [a,b]$ be such that $\gamma\big([c,d]\big)$ is contained in the domain of the local chart; we still denote by λ_0 the restriction of λ_0 to $[c,d]$.

Since γ is a critical point of $\widetilde{\mathcal{L}} - \lambda\circ\Theta$, by standard computations it follows that the following equality holds:

$$\begin{aligned}(4.14)\qquad \int_c^d \frac{\partial\widetilde{L}}{\partial q}\big(t,q(t),\dot q(t)\big)\,v(t) + \frac{\partial\widetilde{L}}{\partial \dot q}\big(t,q(t),\dot q(t)\big)\,\dot v(t)\,\mathrm{d}t&\\ - \Big\langle \lambda_0,\, \frac{\partial\theta}{\partial q}\Big|_{(t,q(t))}\big(v(t),\dot q(t)\big) + \theta_{(t,q(t))}\,\dot v(t)\Big\rangle &= 0,\end{aligned}$$

for every vector field v of class C^1 along γ having support in $]c,d[$; in the formula above we have regarded the derivative $\frac{\partial\theta}{\partial q}\big|_{(t,q(t))}$ as an $\mathbb{R}^{n-k}$-valued bilinear map in $\mathbb{R}^n$. In terms of the local coordinates, the maps θ, $\frac{\partial\theta}{\partial q}(\cdot,\dot q)$, $\frac{\partial\widetilde{L}}{\partial q}$ and $\frac{\partial\widetilde{L}}{\partial\dot q}$ evaluated along γ will be interpreted as follows:

- $\theta \in C^1\big([c,d],\mathrm{Lin}(\mathbb{R}^n,\mathbb{R}^{n-k})\big)$;
- $\dfrac{\partial\theta}{\partial q}(\cdot,\dot q) \in C^0\big([c,d],\mathrm{Lin}(\mathbb{R}^n,\mathbb{R}^{n-k})\big)$;
- $\dfrac{\partial\widetilde{L}}{\partial q},\ \dfrac{\partial\widetilde{L}}{\partial\dot q} \in C^0\big([c,d],\mathbb{R}^{n*}\big)$,

where $\mathrm{Lin}(\cdot,\cdot)$ denotes the space of linear maps between two given vector spaces.

Using the definition of derivative for generalized functions, from (4.14) we get:

$$(4.15)\qquad \Big\langle -\Big(\frac{\partial\widetilde{L}}{\partial\dot q}\Big)' + \frac{\partial\widetilde{L}}{\partial q} - \lambda_0\,\frac{\partial\theta}{\partial q}(\cdot,\dot q) + (\lambda_0\,\theta)', v\Big\rangle = 0,$$

for every C^1 map $v : [c,d] \to \mathbb{R}^n$ having support in $]c,d[$, and, by density, for every $v \in C^1_0\big([c,d],\mathbb{R}^n\big)$. It follows:

$$(4.16)\qquad -\Big(\frac{\partial\widetilde{L}}{\partial\dot q}\Big)' + \frac{\partial\widetilde{L}}{\partial q} - \lambda_0\,\frac{\partial\theta}{\partial q}(\cdot,\dot q) + \lambda_0'\,\theta + \lambda_0\,\theta' = 0.$$

Let $X_1,\ldots,X_{n-k}$ be a referential of $\mathcal{D}'$ along γ; in terms of the local coordinates the X_i's will be thought as elements of $C^1\big([c,d],\mathbb{R}^n\big)$. Moreover, we set

$$X = (X_1,\ldots,X_{n-k}) \in C^1\big([c,d],\mathrm{Lin}(\mathbb{R}^{n-k},\mathbb{R}^n)\big),$$

where the $(n-k)$-tuple $\big(X_1(t),\ldots,X_{n-k}(t)\big)$ is identified with the linear map that takes the i-th vector of the canonical basis of $\mathbb{R}^{n-k}$ to $X_i(t)$.

Composing (4.16) with X, we obtain:

$$\lambda_0'\,\theta(X)+\lambda_0\,\theta'(X)-\lambda_0\,\frac{\partial\theta}{\partial q}(X,\dot q)+\frac{\partial\widetilde L}{\partial q}\,X-\Big(\frac{\partial\widetilde L}{\partial\dot q}\Big)'\,X=0. \tag{4.17}$$

Evaluating (4.8) at X_i with $v=\gamma'$ and using the horizontality of γ we get:

$$\frac{\partial\widetilde L}{\partial\dot q}\,X_i=0,\quad i=1,\ldots,n-k; \tag{4.18}$$

hence:

$$\Big(\frac{\partial\widetilde L}{\partial\dot q}\Big)'X=-\frac{\partial\widetilde L}{\partial\dot q}\,X'\in C^0\big([c,d],(I\!R^{n-k})^*\big). \tag{4.19}$$

Now, considering that $\theta(X)\in\mathrm{Lin}\big(I\!R^{n-k},I\!R^{n-k}\big)$ is invertible, by (4.19) we can write (4.16) in the form:

$$\lambda_0'=\lambda_0\,h_1+h_2, \tag{4.20}$$

with $h_1\in C^0\big([c,d],\mathrm{Lin}(I\!R^{n-k},I\!R^{n-k})\big)$ and $h_2\in C^0\big([c,d],(I\!R^{n-k})^*\big)$.

Applying three times Corollary 4.5, from (4.20) we conclude that λ_0 belongs to the space $C^1\big([c,d],(I\!R^{n-k})^*\big)$; now Lemma 4.6 implies that $\lambda_0\in C^1\big([a,b],(I\!R^{n-k})^*\big)$.

By Lemma 4.7, there exist $\sigma_a,\sigma_b\in(I\!R^{n-k})^*$ such that:

$$\lambda(\alpha)=\int_a^b\lambda_0\,\alpha\,\mathrm{d}t+\sigma_a\,\alpha(a)+\sigma_b\,\alpha(b),\quad \alpha\in C^0\big([a,b],I\!R^{n-k}\big). \tag{4.21}$$

To conclude the proof we show that $\sigma_a=\sigma_b=0$. Let's show for instance that $\sigma_a=0$; the proof of the equality $\sigma_b=0$ is analogous.

Using local charts around $\gamma\big([a,d]\big)$, for d close to a, we consider variational vector fields v of class C^1 supported in $[a,d[$, with $v(a)\in T_{\gamma(a)}P$. Arguing as in the deduction of formula (4.14), we get the following equality:

$$\begin{aligned}\int_a^d\frac{\partial\widetilde L}{\partial q}\big(t,q(t),\dot q(t)\big)\,v(t)+\frac{\partial\widetilde L}{\partial\dot q}\big(t,q(t),\dot q(t)\big)\,\dot v(t)\,\mathrm{d}t\\ -\int_a^d\lambda_0(t)\Big[\,\frac{\partial\theta}{\partial q}\Big|_{(t,q(t))}\big(v(t),\dot q(t)\big)+\theta_{(t,q(t))}\,\dot v(t)\Big]\,\mathrm{d}t\\ -\sigma_a\Big[\,\frac{\partial\theta}{\partial q}\Big|_{(a,q(a))}\big(v(a),\dot q(a)\big)+\theta_{(a,q(a))}\,\dot v(a)\Big]=0.\end{aligned} \tag{4.22}$$

From Corollary 4.5 and formula (4.16) it follows that $\frac{\partial\widetilde L}{\partial\dot q}$ is of class C^1, and we can thus use integration by parts in (4.22) to obtain an equality of the form:

$$\int_a^d u(t)\,v(t)\,\mathrm{d}t+\sigma_a\,\theta_{(a,q(a))}\,\dot v(a)=0, \tag{4.23}$$

for some $u\in C^0\big([a,d],I\!R^{n*}\big)$, whenever v is chosen with $v(a)=0$. By considering arbitrary v supported in $]a,d[$, from (4.23) we obtain that $u\equiv0$ in $[a,d]$, so that the integral in (4.23) vanishes for all v. Now, we can choose v with $v(a)=0$ and $\dot v(a)$ arbitrary, so that (4.23) implies that $\sigma_a=0$, because $\theta_{(a,q(a))}$ is surjective. This concludes the proof. $\square$

REFERENCES

[1] R. Abraham, J. E. Marsden, *Foundations of Mechanics*, 2nd Edition, Benjamin/Cummings, Advanced Book Program, Reading, Mass., 1978.

[2] V. I. Arnold, V. V. Kozlov, A. I. Neishtadt, *Mathematical Aspects of Classical and Celestial Mechanics*, Dynamical Systems III, Encyclopaedia of Mathematical Sciences, Ed. V. I. Arnold, Chapter I (1988), 1–286.

[3] J.-M. Bismut, *Large Deviations and the Malliavin Calculus*, Progress in Mathematics, Birkhäuser, Boston - Basel - Stuttgart, 1984.

[4] R. L. Bryant, L. Hsu, *Rigidity of Integral Curves of Rank* 2 *Distributions*, Invent. Math. **114** (1993), 435–461.

[5] F. Cardin, M. Favretti, *On Nonholonomic and Vaconomic Dynamics of Mechanical Systems with Nonintegrable Constraints*, J. Geom. Phys. **18** (1996), 295–325.

[6] M. do Carmo, *Riemannian Geometry*, Birkhäuser, Boston, 1992.

[7] I. Kishimoto, *The Morse index theorem for Carnot-Carathéodory spaces*, J. Math. Kyoto Univ. **38** (1998), no. 2, 287–293.

[8] I. Kupka, W. M. Oliva, *The Non-Holonomic Mechanics*, Journal of Differential Equations **169** (2001), no. 1, 169–189.

[9] S. Lang, *Differential Manifolds*, Springer-Verlag, Berlin, 1985.

[10] W. Liu, H. J. Sussmann, *Shortest Paths for Sub-Riemannian Metrics on Rank–*2 *Distribution*, Memoirs AMS **564**, vol. 118, 1995.

[11] R. Montgomery, *A Survey of Singular Curves in Sub-Riemannian Geometry*, J. Dyn. Contr. Sys. **1** (1995), 49–90.

[12] R. Montgomery, *Abnormal Minimizers*, SIAM J. Contr. Optim. **32**, vol. 6 (1994), 1605–1620.

[13] R. Montgomery, *Singular Extremals on Lie Groups*, Math. Control Signal Systems **7** (1994), 217–234.

[14] J. R. Munkres, *Elements of Algebraic Topology*, The Benjamin/Cummings Publishing Company, Inc., 1984.

[15] R. Palais, *Foundations of Global Nonlinear Analysis*, W. A. Benjamin, 1968.

[16] P. Piccione, D. V. Tausk, *On the Banach Differentiable Structure for Sets of Maps on Non Compact Domains*, Nonlinear Analysis: Series A, Theory and Methods, vol. 2 (**46**), pgs. 245–265 (2001).

[17] P. Piccione, D. V. Tausk, *Variational Aspects of the Geodesics Problem in sub-Riemannian Geometry*, Journal of Geometry and Physics **39** (2001), 183–206.

[18] P. Piccione, D. V. Tausk, *On the Maslov and the Morse Index for Constrained Variational Problems*, to appear in the Journal de Mathématiques Pures e Appliquées (2001).

[19] A. M. Vershik, V. Y. Gershkovich, *Nonholonomic Dynamical Systems, Geometry of Distributions and Variational Problems*, Dynamical Systems VII, Encyclopaedia of Mathematical Sciences, Eds. V. I. Arnold and S. P. Novikov (1990), 4–79.

DEPARTAMENTO DE MATEMÁTICA,
UNIVERSIDADE DE SÃO PAULO, BRAZIL
E-mail address: piccione@ime.usp.br
URL: http://www.ime.usp.br/˜piccione

DEPARTAMENTO DE MATEMÁTICA,
UNIVERSIDADE DE SÃO PAULO, BRAZIL
E-mail address: tausk@ime.usp.br
URL: http://www.ime.usp.br/˜tausk